Psychological Harms of Screening for Type 2 Diabetes Mellitus: A Systematic Review

By

Russell Coletti

A Master's Paper submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Public Health in the Public Health Leadership Program

Chapel Hill

2013

Russell Harris, MD MPH / Advisor

Date

Colleen Barclay, MPH / Second Reader

Date

Table of Contents

Abstract

Background: Recommendations for screening for Type 2 Diabetes Mellitus (T2DM) are based on evidence containing large amounts of uncertainty regarding the relative benefits and harms of screening. Evidence is particularly scarce about the potential psychological harms.

Objective: This systematic review examines the literature to determine what is currently known about the psychological harms of screening for T2DM.

Methods: PubMed, EMBASE, PsycINFO, and reference lists of published literature were searched for English-language studies published up until April 5th, 2013. Studies assessing psychological outcomes for at least one screened group within 3 years of screening were included. Data from included studies were extracted and strength of evidence was assessed for each psychological outcome.

Results: Out of 4,435 search results, 14 articles were included. Screening produces a dose-dependent effect on levels of anxiety and worry shortly after testing. A diagnosis of screen-detected diabetes is associated with higher cognitive distress as well as thought intrusiveness. Depression appears to be associated with treatment of screen-detected diabetes, but not the actual screening test or diagnosis itself. Screening appears to have no effect on general psychological quality of life or negative well-being. Overall strength of evidence is low due to the few number fair of studies assessing each outcome.

Conclusion: Overall, the available literature is insufficient to say with high certainty that screening for T2DM results in meaningful psychological harm, though the included studies hint that screening is not entirely benign.

Acknowledgements

Thank you to the following people for your help in producing this paper:

Russell Harris

Colleen Barclay

Mellanye Lackey

Jamie Carter

Introduction

At first glance, type 2 diabetes (T2DM) appears to represent a reasonable target for screening. It affects a large percentage of the population, with an estimated 25 million people, or 8.3% of the U.S. population having diabetes.[1] Of those 25 million people, 27% remain undiagnosed, largely because diabetes has an identifiable latent phase, meaning people often have blood sugar levels consistent with diabetes for years without displaying any signs or symptoms.[2,3] Another 79 million people meet criteria for "pre-diabetes," (also called impaired glucose tolerance and impaired fasting glucose) a term used to classify the people with blood glucose higher than the "normal" range but not high enough to meet the threshold for diabetes.[1,3] Additionally, valid and reliable tests can detect diabetes in asymptomatic individuals.[2] Therefore, screening has the potential to identify a large number of people with asymptomatic T2DM, allowing for early detection and treatment to reduce negative health outcomes.

In addition to the above qualities, however, a screening program must demonstrate a net benefit large enough to justify its implementation.[4] Unfortunately, evidence about potential benefits and harms of T2DM screening is scarce. For example, one of the most important screening tenets dictates that treating a disease early (in the asymptomatic phase) must produce a net morbidity and/or mortality benefit compared to treating only clinically-apparent disease.[4] Evidence up to this point looking at T2DM screening contains a large degree of uncertainty on this issue.

Ideally, we would like a randomized controlled trial (RCT) that assesses the direct effect of screening for T2DM on health outcomes. Unfortunately, only one such study has been done. The ADDITION-Cambridge study was a cluster randomized screening trial that was nested

within a larger trial looking at intensive vs. usual treatment of screen-detected diabetes. After a 10 year follow-up, the study found no difference in all-cause or cardiovascular-related mortality between screened and unscreened groups.[5] Aside from this limited trial data, a 2008 systematic review performed for the U.S. Preventive Services Task Force (USPSTF) found mixed results about potential benefits from only a small collection of observational and modeling studies.[2]

Despite very low certainty about the benefits of screening thus far, many guideline development groups recommend some form of screening, generally among groups they deem to be higher risk for diabetes. **Table 1** lists several such groups and their recommendations. Of the groups that report grades along with their guidelines, many indicate that they based these recommendations on moderate to low quality evidence, meaning that as new evidence emerges, these recommendations may change.

Despite no definitive evidence of substantial benefit, these groups still recommend routine screening likely because there may be benefit for some people while there appears to be little demonstrable harm.[2,3] The screening cascade generally contains multiple parts. First, patients undergo initial blood tests, which if positive, will likely result in a confirmatory test. If a person is diagnosed, then he or she may begin a variety of treatments aimed at lowering blood sugar. The person may or may not subsequently progress to symptomatic disease, which can, in turn, produce negative health outcomes in the future. At any point along this cascade, the possibility of harm exists.

Table 1. List of guideline development groups and their recommendations. Strengths of the recommendations (if applicable) are as stated by their respective group and may be different from those of other groups.

Group	Recommendations (***Strength***)

American Diabetes Association[6]	- Recommends screening adults of any age who are obese and have one or more of the following *(B: supportive evidence from well-conducted cohort studies)*:
	  - Physical inactivity
	  - First degree relative with diabetes
	  - High-risk ethnicity
	  - Women with history of gestational diabetes or delivered a baby weighing >9 lbs
	  - Hypertension (BP 140/90) or on therapy
	  - HDL <35 mg/dL and/or triglycerides >250 mg/dL
	  - Women with polycystic ovarian syndrome
	  - A1C >5.7%, IGT or IFG on previous testing
	  - Other clinical conditions associated with insulin resistance (severe obesity, acanthosis nigricans)
	  - History of cardiovascular disease
	- In the absence of these factors, screen beginning at 45 years *(E: expert consensus)*
	- If "prediabetes," test yearly

The Endocrine Society[7]	- Recommends screening with fasting plasma glucose as part of screening for the components of metabolic syndrome. *(strong recommendation: moderate quality evidence)* - This should be done every 3 years *(weak recommendation: very low quality evidence)* - Recommends screening for type 2 diabetes every 1-2 years in people previously diagnosed with IFG or IGT *(weak recommendation: very low quality evidence)*

American Heart Association[8]	- Recommends screening in adults who are:
	  - Overweight and over age 45
	  - Overweight, under age 45 and one or more of the following
	  - Hypertension
	  - High cholesterol
	  - Family history of diabetes
	  - African-American, Asian-American, Latino, Native American or Pacific Islander
	  - History of gestational diabetes or delivered a baby > 9 lbs
	- Recommends screening every 3 years if tests are normal
	- Recommends screening every 1-2 years if prediabetes

United States Preventive Services Task Force[9]	- Recommends screening in adults with sustained blood pressure >135/80 *(B:high certainty that net benefit is moderate or there is moderate certainty that net benefit is moderate to substantial)*
National Institute for Health and Care Excellence[10]	- Recommends screening for those with a high risk score on a validated risk-assessment tool or self-assessment questionnaire. - Recommends screening for those aged 25 and over of South Asian or Chinese origin who have a BMI > 23 kg/m^2
Canadian Task Force on Preventive Health Care[11]	- Recommends screening adults at high risk of diabetes (33% chance of developing diabetes within 10 years as determined by validated risk score) every 3-5 years. *(weak recommendation: low-quality evidence)* - Recommends screening adults at very high risk (50%) annually *(weak recommendation: low-quality evidence)* - Recommends not screening those at low to moderate risk (1-17%) *(weak recommendation: low-quality evidence)*
International Diabetes Federation[12]	- Recommends screening those at "high risk." No interval specified - Recommends identifying those at "high risk" using questionnaire to assess age, waist circumference, family history, cardiovascular history, and gestational history.

In general, the potential harms of screening for T2DM can be divided into two categories: physical and psychological. Physical harms largely include possible consequences of treating screen-detected diabetes, such as hypoglycemia, cardiovascular outcomes, and more minor drug-specific side effects.[2] Psychological harms are broader, including effects from an invitation to screening, the actual screening test, false positives, false negatives, labeling, and distress related to disease management. Aside from the potential physical and psychological harms present throughout the cascade, the patient must also endure the hassle of spending time and money undergoing tests and receiving treatment.

Unfortunately, evidence about the effects of screening on psychological outcomes is lacking, as diabetes-related studies tend to focus more on "harder" clinical outcomes.[2,3,13] Therefore, recommendations for T2DM screening actually incorporate incomplete harms

evidence, meaning that screening large numbers of individuals could, in fact, be causing harm. It is therefore important to understand what is currently known and unknown about the potential harms of screening for T2DM. The objective of this review is to systematically examine the current evidence for the harms of screening for T2DM. For the purposes of this review we will focus only on psychological harms.

Methods

This systematic review examines the evidence about the psychological harms of screening for T2DM. In this review, I defined screening as the testing of asymptomatic individuals to detect "abnormal" blood glucose levels. These tests include blood tests for Hemoglobin A1C (A1C), Fasting Plasma Glucose (FPG), and Oral Glucose Tolerance Test (OGTT).

Although the screening cascade contains many facets, for simplicity I have divided the screening process into 2 phases: (1) the process of the actual screening test, subsequent work-up, and labeling of the patient; and (2) the process of treating the screen-detected patient. Two Key Questions (KQs) correspond to these two phases:

Key Question 1 (KQ1): What are the psychological harms of <u>screening</u> for T2DM?

Key Question 2 (KQ2): What are the psychological harms of <u>treating</u> screen-detected T2DM?

Table 2: Eligibility Criteria for Included Studies

	Inclusion
Population	Adults age >18 years who have one or more of the following: -Undergone screening for abnormal blood glucose levels -Screen-detected type 2 diabetes (non-insulin dependent diabetes) usually defined as the combination of any two or more of the following[6]: - Hemoglobin A1C ≥ 6.5% - Fasting Plasma Glucose ≥ 126 mg/dL - 2hr Plasma glucose ≥ 200 mg/dL during an OGTT -Impaired glucose tolerance defined as 2hr plasma glucose 140-200 mg/dL -Impaired fasting glucose defined as fasting level of 100-125 mg/dL -Individuals with asymptomatic type 2 diabetes diagnosed <3 years from measurement of outcome. Individuals must be previously undiagnosed with depression, anxiety, or other psychological disorder prior to screening
Intervention	Screening for type 2 diabetes, impaired glucose tolerance, or impaired fasting glucose using Hemoglobin A1C, Fasting Plasma Glucose, or Oral Glucose Tolerance Test
Comparison	Adults without screening or screen detected diabetes (including people who have undergone screening but have received a negative result)
Outcomes	Psychological outcomes during or after screening including: -Depression -Anxiety -Worry -Fear -Distress -Decrease in psychological quality of life
Timeframe	From first screening to 3 years after diagnosis of diabetes or IGT/IFG
Time Period	1966-present
Study Design	-Randomized Controlled Trials -Cohort Studies -Case-Control Studies -Qualitative Studies -Systematic Reviews
Study setting	Primary care clinics at the beginning of study

Eligibility Criteria

I determined eligibility criteria in the form of PICOTTSS (Population, Intervention, Comparators, Outcomes, Timeframe of study, Time period of search, Study designs accepted, and Study setting accepted). **Table 2** details all of the inclusion criteria.

I included studies of adults 18 years and older, including both men and women, who underwent T2DM screening or have screen-detected T2DM or Impaired Fasting Glucose/Impaired Glucose Tolerance (IFG/IGT). I included only individuals without depression, anxiety or any other psychological disorders prior to screening, as I was interested only in the effect of screening on psychological well-being, not the incidence of T2DM in those with pre-existing mood disorders. To answer KQ2, I included studies investigating psychological well-being among those receiving treatment for screen-detected diabetes. As an indirect measure of screen-detected diabetes, I also included individuals with asymptomatic diabetes of 3 years duration or less. I excluded studies if the study did not explicitly identify one group as having screen-detected or asymptomatic T2DM within the specified time frame. For randomized controlled trials (RCTs), qualitative studies, cohort, and case-control studies, relevant comparators included those who have not undergone screening for T2DM, those who have clinically diagnosed T2DM i.e., received a diagnosis after displaying signs and symptoms, or those with a negative result from a screening test.

I included any study investigating psychological outcomes related to screening including depression, anxiety, worry, fear, distress, diabetes-related distress, or change in psychological quality of life. I considered outcomes within 3 years of diagnosis, because those receiving

standard treatment for screen-detected diabetes may not experience distress until after the first year following diagnosis.[14]

Search Strategy

I searched PubMed®, Embase, and PsycINFO® for all systematic reviews and studies published up until April 5, 2013. I chose to include English-language cohort, case-control and qualitative studies in addition to systematic reviews and randomized trials in anticipation that very few of the latter studies would be available on this topic. I also reviewed the reference lists of included studies for any other relevant citations.

To develop the search strategy, I consulted a health sciences librarian, who assisted in the development of relevant search terms. For interventions, I used the Medical Subject Heading (MeSH term) "diabetes mellitus, type 2" as well as the key words "non insulin dependent diabetes," "impaired fasting glucose" and "impaired glucose tolerance," along with the MeSH terms "mass screening," "hemoglobin A, glycosylated," and "glucose tolerance test," as well as key words such as "screen," "early diagnosis," "early detection," and "fasting glucose." Finally, for the outcomes we used key words such as "depression," "distress," "worry," "fear," "anxiety," and "quality of life." Of note, the term "diabetes-related distress" is a disease-specific outcome referring to psychological effects of carrying the diagnosis of and receiving treatment for diabetes.[15] I did not include "diabetes-related distress" as a separate term, because I expected that the broader term "distress" would be sensitive enough to find any studies examining this diabetes-specific outcome. Due to differences among the databases' index words, search terms were slightly different for each data source. The full search strategy is listed in **Appendix A**.

Data Management

All titles and abstracts were imported into RefWorks 2.0 (ProQuest, LLC). I reviewed and compared the titles and abstracts against inclusion and exclusion criteria, excluding those that clearly did not meet the criteria. I then reviewed the remaining full-text articles to determine which studies would ultimately be included. Due to limited resources a dual-review was not performed; however in the event of a dual-review, if either reviewer considered a result eligible, it would be included for full-text review. Then, the independent reviewers would review the full-text articles, and if a subsequent consensus was not reached, a third, senior reviewer would resolve disagreements about inclusion.

Data from included studies was then abstracted into evidence tables. These data include study type, participant information, intervention, comparators, outcome assessment, and results. Here again, a dual-review was not feasible, but would have included a second reviewer checking data tables for consistency. Disagreements would have been settled by discussion and consensus.

Strength of Evidence

I adapted criteria from the United States Preventive Services Task Force (USPSTF) and the Evidence-based Practice Center program of the Agency for Healthcare Research and Quality to rate strength of evidence.[16,17]

I first rated the internal and external validity of each individual study. Internal validity refers to the risk of bias in each study, while external validity refers to what degree the study results can be generalized to other populations. Risk of bias for each individual study was evaluated according to USPSTF criteria for internal validity.[16] These criteria include creation and preservation of comparable groups, high degree of follow-up, measurement of intervention and

outcomes, and appropriate analysis that addresses and corrects for potential confounders, including an intention to treat analysis for RCTs. Based on these criteria, each study received a rating of good, fair, or poor. Good studies met all of the listed criteria and had the least risk for bias. Fair studies generally met all criteria without any major flaws, though still were susceptible to some bias. Poor studies had at least one major flaw such as non-comparable study groups, severe attrition rates, or inattention to key confounders. Poor studies were excluded from further review.

I evaluated the generalizability for each study using the USPSTF criteria for external validity.[16] These criteria call for examining each study's population, situation, and providers, each of which may contain characteristics limiting the results' generalizability. For example, a study population may include or exclude certain ages, genders, ethnicities, comorbidities, or income levels that affect the outcomes of interest. The study situation is the degree to which the environment or context would likely be replicated, for instance whether the intervention of interest is as readily available among general practice as it is in the study. Finally, the providers in the study may have different training or skills than the majority of providers nationwide. Depending on the degree of generalizability, each study received a rating of good, fair, or poor. Good studies differed slightly from general population, situation, and providers such that the study results still had a high likelihood of occurring in usual practice. Fair studies contained differences that moderately affected likelihood of attaining results in general practice, while poor studies contained substantial differences with low likelihood of applying to general practice. I did not exclude studies with poor generalizability, as they could still have internally valid results that apply to a specific population or setting.

Next I rated aggregate internal and external validity of the studies for each outcome. Aggregate validity not only synthesizes individual study validities; it also addresses the strength of each study design and whether the study designs were sufficient to answer the question. I gave a rating of high, moderate, or low for both overall internal and external validity.

I also rated the consistency of evidence for each outcome. Consistency refers to the degree to which the evidence shows a uniform direction and magnitude of effect.[17] For my results, a study could show strongly positive, mildly positive, negative, or no difference in psychological outcomes. Generally evidence is considered inconsistent if results contain both large positive and negative effects; however I also considered results inconsistent if they showed any degree of no difference in addition to large positive or negative effects. If multiple studies were included for a key question, I gave a rating of either "consistent" or "inconsistent." If there was only one study for a given key question, I gave a rating of "not applicable."[17]

Finally, I rated precision for each outcome. Precision is the statistical and conceptual certainty of an intervention's effects.[16,17] Statistical certainty includes the confidence interval around a point estimate. Conceptual certainty considers the number and size of the studies that form the evidence base. I gave ratings of "high," "moderate," or "low," along with justifications for each rating. Evidence with high precision had an adequate number of well-performed large studies that showed consistent results. Moderately precise evidence may have some degree of uncertainty due to small inconsistencies, fewer studies, or small faults in study quality. Evidence with low precision suffers from limitations such that one cannot draw conclusions due to large statistical uncertainty, very few studies or markedly inconsistent results. Evidence with moderate or low precision is likely to change in magnitude and/or certainty with the development of new evidence.

Data Synthesis

I qualitatively synthesized data to answer each key question using their outcomes and strength of evidence. Then, using the evidence from each key question, I drew an overall conclusion about the psychological harms associated with screening for type 2 diabetes. I did not perform a meta-analysis due to significant heterogeneity in outcomes.

Results

I identified a total of 14 articles related to 8 studies for inclusion. An electronic search of PubMed®, Embase, and PsycINFO®, produced a total of 4,429 results, and a review of references produced another 6 results. After a removal of duplicates, 3,936 results remained. I excluded 3,866 results after a title and abstract review, leaving 70 articles for full-text review. Of these 70, 56 articles were excluded, the majority for wrong publication type; see flow diagram **Figure 1.** Many of the other studies excluded during full-text review did not have appropriate comparison groups or did not specify that diabetes was screen-detected.

Characteristics of Included Studies

Fourteen results remained for inclusion: 11 prospective cohort studies, 1 prospective qualitative study, 1 randomized controlled trial, and 1 non-randomized controlled trial. The sample sizes ranged from 23 to 7,380 for a total of 19,860 participants, and the mean age of the participants ranged from 55-68 years. Study duration ranged from 6 weeks to 13 years. Eight studies took place in the U.K., 3 in The Netherlands, 2 in the U.S., and 1 in Germany. Individual study characteristics can be found in **Appendix B**.

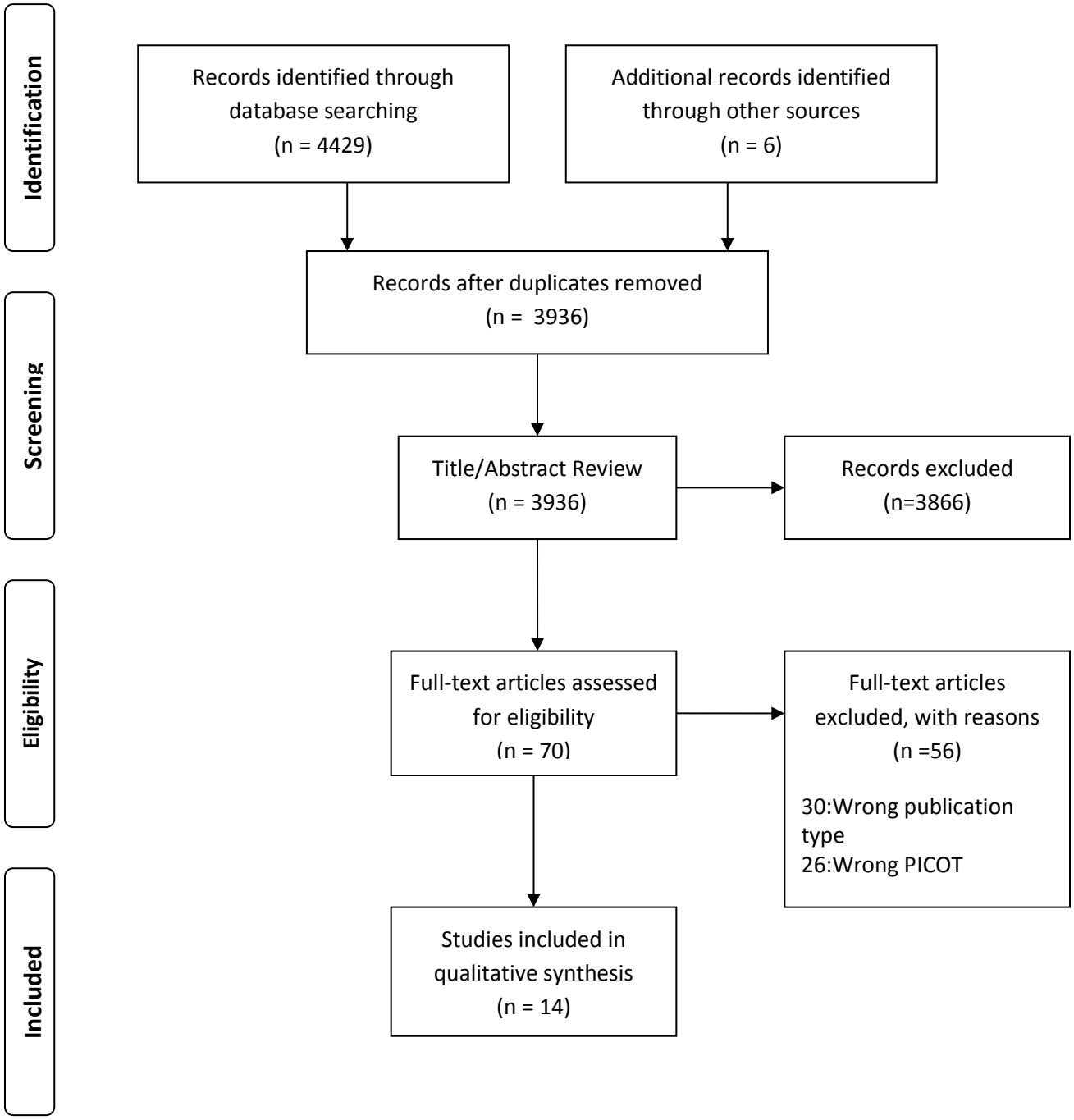

Figure 1. Flow diagram of included studies (Adapted from the PRISMA statement[18])

All included studies had at least 1 group that underwent screening for diabetes. Three studies compared screened and unscreened groups;[19-21] two compared outcomes between people with screen-detected diabetes and people with diabetes detected in general practice;[22,23] six studies compared those who screened positive with those who screened negative for T2DM;[24-29] and two assessed outcomes between groups undergoing screening several years apart.[30,31] Finally, the qualitative study investigated the experiences of people undergoing a stepwise screening process.[32]

Most studies employed targeted screening on "high-risk" populations. Those associated with the ADDITION trial, a randomized trial in which primary care clinics were cluster randomized to screening or no screening, identified individuals using a risk factor questionnaire.[19-21,29,32] Studies associated with the Hoorn Screening Study, a targeted screening project carried out in a Dutch population, used a similar diabetes symptom and risk questionnaire.[22,24,25] One study screened siblings of people with T2DM.[27] Five studies did not specifically target higher risk individuals; however two of these drew study populations from larger studies investigating correlates of cardiovascular disease.[23,28]

The most commonly assessed outcome was general mental health(five studies),[22,24,26,30,31] followed by anxiety (four studies)[20,21,27,29] and depression(four studies).[21,23,28,29] Three studies reported changes in negative well-being,[22,24,25] and two of these also measured cognitive distress.[22,25] Worry,[21] false reassurance,[19] and frequency of thought about what it would be like to have diabetes[27] were each measured once. The tools used to measure the various psychological outcomes are listed in **Table 3**.

Table 3. Characteristics of Psychological Outcomes Assessment Tools

Outcome	**Assessment Tool**	**Description**
Mental Health	Medical Outcomes Survey Short Form 36 item (SF-36)	A generic quality of life measurement tool with a physical subscale and a mental subscale[33] Higher scores indicate better quality of life
Anxiety	1. Spielberger State Anxiety Inventory Short Form (SSAI-SF) 2. Hospital Anxiety and Depression Scale-Anxiety (HADS-A)	1. A generic measurement tool that assesses state of short-term anxiety[34] >40 indicates clinically significant anxiety 2. A generic measurement tool that assesses generalized anxiety 0-7 No anxiety 8-10 Mild anxiety 11-14 Moderate anxiety 15-21 Severe anxiety
Negative Well-Being "Negative mood"[25]	Negative Well Being Subscale of Dutch Short Version of the Well-Being Questionnaire (WBQ-12)	Generic quality of life measurement tool that assesses depression and anxiety >4 indicates elevated depressive symptomatology
Depression	1.Hospital Anxiety and Depression Scale-Depression (HADS-D) 2.Center for Epidemiologic Studies Depression Scale (CES-D)	1. Generic measurement tool that assesses depressive symptoms.[35] 0-7 No depression 8-10 Mild 11-15 Moderate >15 Severe 2. Generic measurement tool assessing current level of depressive symptoms >16 indicates clinical depression
Cognitive Distress	Diabetes Symptom Checklist	A diabetes-specific checklist assessing occurrence of symptoms over eight domains, one of which is cognitive distress[25]

Outcome	Assessment Tool	Description
False Reassurance	2 questions asked to patients to indicate: 1. Perceived risk 2. Comparative risk (non-validated measure)	1. Perceived risk: chance of getting diabetes in the future; expressed as a percentage 0-100% on 11-point scale[19] 2. Comparative risk: chance of getting diabetes compared to other people of their own age; 1: much lower to 5: much higher
Worry	6 item worry scale (non-validated measure)	Assesses worry about developing diabetes in the future. Adapted from Lerman cancer worry scale[21,36]
Measures of Cognition (knowledge and beliefs)	3 questions that elicit thoughts about diabetes (non-validated measure)	1. Perceived risk: likelihood of developing diabetes; expressed as Likert scale "very likely" to "not at all likely"[27] 2. Occupied thoughts: how much the person thinks about what it would be like to have diabetes; "never" to "all the time" 3. Thoughts on complications: likelihood of developing complications if diagnosed with diabetes; "very unlikely" to "very likely"

Most of the tools used to measure psychological outcomes were generic tools that are generally used for people with various conditions. Only four of the ten tools were specific to diabetes: the Diabetes Symptom Checklist, and the study-specific questions assessing false reassurance, worry, and knowledge and beliefs about diabetes. Out of those four, only one, the Diabetes Symptom Checklist, appeared to be used and validated in other studies.[22,24,25]

Key Question 1 (KQ1): What are the psychological harms of screening for T2DM?

Mental Health

Five studies measured mental health; none of these compared screened and unscreened populations. Compared to people with diabetes detected in general clinical practice, those with

screen-detected diabetes had significantly better SF-36 Mental Health Component (MHC) scores from 2 weeks to 1 year of follow-up (mean scores at 1 year 54.3 vs. 50.8; p=0.009).[22] However, when adjusted for multiple comparisons, this difference is no longer significant. Among people screened for diabetes, there was no significant difference in scores at 1 year between those who screened positive and those who screened negative.[24,26] Additionally, among those screened, there was no difference in mental health scores between people 1-3 years after screening and people 10-13 years after screening.[30,31]

Anxiety

Four studies measured anxiety, two of which were trials comparing screened and unscreened groups. In one, screened individuals had similar scores on SSAI and HADS-A as unscreened individuals at the time of screening and at 3-6 and 12-15 months.[21] Neither of the groups had mean scores suggesting clinically important anxiety at any time. In contrast, the second trial showed that those invited for screening had significantly higher mean anxiety scores at 6 weeks than those not screened (37.6 vs. 34.1, respectively; p=0.015); although, neither of these scores represents clinically important anxiety.[20] Of note, a subgroup analysis indicated an increasing trend in anxiety scores corresponding with progression through the step-wise screening process, with a significant difference between those ultimately diagnosed with diabetes and those testing negative at initial screening. Furthermore, the mean anxiety score among those diagnosed indicated clinically important anxiety.[20]

In a screening study of siblings of people with diabetes, there was no difference in anxiety scores between those testing positive and those testing negative at 1 year follow-up.[27] However, both groups had significantly lower scores at 1 year compared to "baseline" values

right before screening. A similar study showed no difference in the percentage of people meeting criteria for anxiety at 1 year between those with positive test results and those with negative test results.[29]

Depression

Four studies assessed depressive symptoms; only one of these studies compared screened and unscreened groups. In this trial, there was no difference in HADS Depression scores between the screened and unscreened groups at 3 and 12 months after screening.[21] Studies comparing screen-positive and screen-negative groups showed no difference in depression scores at 1 year.[27,29] One study compared a group with screening-confirmed normoglycemia to a group with impaired fasting glucose and one with untreated diabetes.[28] Neither of the latter groups had an increased odds of depressive symptoms over 5 years compared to the normoglycemic group.

Negative Well-being

Three related studies measured negative well-being, which they equated to negative mood. Compared to people with diabetes newly detected in clinical practice, those with screen-detected diabetes had similar scores on the Negative Well Being subscale of the Well-Being Questionnaire.[22] There was also no difference between individuals with screen-detected diabetes and those that screened negative.[22,24,25]

Cognitive Distress

Two of the three studies that measured negative well-being also assessed cognitive distress as one domain in an array of diabetes symptoms; however only one of the studies

compared differences between groups;[25] the other assessed for change in distress over time within groups.[22] In the latter study, people with diabetes detected in general practice had significantly lower distress at 1 year as compared to 2 weeks after diagnosis, whereas those with screen-detected diabetes did not change significantly over the same period.[22] At 1 year, the two groups appeared to have similar levels of distress. The second study showed that, compared to people who screened negative for diabetes, a significantly higher percentage of those with screen-detected diabetes had cognitive distress at 2 weeks and 6 months (2 weeks: 64% vs. 49%, p=0.043; 6 months: 64% vs. 45%, p=0.004).[25] At 12 months, 60% of individuals with screen-detected T2DM had reported symptoms of cognitive distress, compared to 52% of those who screened negative; however this difference was no longer significant.

False Reassurance

One study assessed false reassurance between people who screened negative for diabetes and those who were unscreened.[19] False reassurance refers to the idea that after receiving screening test results indicating no disease, a person will mistakenly underestimate their future risk of developing the disease.[19] Compared to unscreened individuals, those receiving a negative screening test estimated a similar chance of developing diabetes in the future. However, immediately after screening those with a negative screening test rated their *comparative* risk significantly lower than did unscreened individuals. In other words, those who screened negative for T2DM were slightly more likely to say that, compared to other people their age, their future risk of developing diabetes was lower. This difference between groups was no longer significant at 6 months and 1 year.

Worry

One study assessed worry about developing diabetes in the future. People who underwent screening for diabetes expressed similar worry as did unscreened controls.[21] However, among those screened, people with a positive result at the initial random blood glucose test showed significantly higher worry about developing diabetes in the future than did those with a negative result (8.18 vs. 7.97, p=0.002). Additionally, people further along in the screening process had significantly higher worry than did those for whom diabetes was ruled out early on in the process. For example, participants that progressed to an oral glucose tolerance test had significantly greater worry about developing diabetes than did those who initially screened negative at the random blood glucose test. Despite these differences between groups, a qualitative study suggested that overall worry remained low, possibly because patients regard initial screening tests as routine and generally expect negative results.[32] Then, when testing positive early in the screening process, patients downplay the importance of the result, attributing it to a recent meal or normal fluctuation. Even when ultimately diagnosed with diabetes, most participants did not express much worry.

Measures of Cognition

One study assessed the degree to which diabetes occupied participants' thoughts.[27] Compared to those with normoglycemia, people whom screening identified as "at risk," were significantly more likely to think about what it would be like to have diabetes (24.2% vs. 13.1%, p=0.006). They were also significantly more likely to think that developing diabetes was "quite or very likely" (58.2% vs. 31.5%, p=0.006). However, those deemed "at risk" were no more likely to expect complications in the event of developing diabetes.

Key Question 2: What are the psychological harms of treating screen-detected T2DM?

There were no studies specifically looking at whether treating screen-detected T2DM produced psychological harms compared to delaying treatment until clinical detection. However, one study compared the development of depressive symptoms among various screened groups stratified by blood glucose levels.[28] Although impaired fasting glucose and untreated diabetes were not associated with depression, people undergoing treatment for screen-detected diabetes had greater odds of having depressive symptoms at 5 years than did those with normal blood sugar (OR 1.52, 95% CI 1.09-2.12).

Synthesis

Due to high levels of heterogeneity in outcomes, follow-up, and study design, the results were qualitatively synthesized. Results for the strength of evidence for each outcome are shown in **Table 4**. In general, strength of evidence for each outcome was low, due to the low number of corresponding studies. Additionally, all of the studies were of fair quality, as significant loss to follow-up was a common problem. Other problems included failure to account for both baseline differences and missing data in the final analysis. The factors limiting external validity included the participant demographics and the screening process itself. For example, many of the studies comprised participants who were at a higher risk than the average screened individual for having undiagnosed diabetes. Additionally, a considerable number of studies used a screening process that had more steps than would the U.S. general practice. For example, participants in the ADDITION trial underwent random glucose, fasting glucose, and oral glucose tolerance tests (progression to a subsequent test depended on prior positive results).[21] Conversely, general diabetes screening the U.S. general practice usually involves 1-2 tests.[6] Such a difference could

influence participants' attitudes toward testing and subsequent diagnosis. Full ratings for internal and external validity are included in **Appendix C** and **Appendix D**, respectively.

Discussion

The purpose of this systematic review was to assess the evidence on the psychological harms of screening for type 2 diabetes. The considerable heterogeneity in comparison groups, outcomes, and results, as well as the overall low strength of evidence as assessed in this review, suggests that the evidence is currently insufficient to say with high certainty that screening for T2DM results in meaningful psychological harm.

Two studies suggest that general psychological well-being remains unaffected among screened individuals, both those with positive and negative results. Additionally, those diagnosed clinically have slightly lower psychological quality of life than those who are screen-detected.[22] However, the strength of evidence for the relationship between psychological well-being and screening is low.

Table 4. **Strength of Evidence for all outcomes stratified by comparison groups.**

Number of studies; # of participants	Design(s)/Overall Internal Validity (reason)	Overall External Validity	Consistency	Precision	Strength of Evidence
KQ1					
Mental Health: Screen-detected vs. General Clinical Practice					
1; 165	Prospective Cohort/ Fair: (high attrition rates; baseline dissimilarities not addressed in analysis)	Fair: (Primarily Caucasian, identified through symptom questionnaire, stepwise screening process)	N/A	Precise	Low
Mental Health: Screen Positive vs. Screen Negative					
2; 1,512	Prospective Cohort/ Fair (Baseline characteristics different/unclear and not addressed in analysis)	Fair to Good: (Primarily Caucasian, identified through symptom questionnaire, one study had a stepwise screening process)	Consistent	Imprecise	Low
Mental Health: 1-3 years post screening vs. 10-13 years post screening					
2; 1,593	Prospective Cohort/ Fair (high attrition rates)	Fair: (Primarily Caucasian, relatively healthy population)	Consistent	Imprecise	Low
Anxiety: Screened vs. Unscreened					
2; 7,735	RCTs/ Fair (high attrition rates)	Fair: (Study population very high risk, maybe more so than average screening population)	Inconsistent	Imprecise	Low
Anxiety: Screen Positive vs. Screen Negative					
2; 3,671	Prospective Cohorts/ Fair: (High attrition rates, unadjusted analysis)	Fair: (Primarily Caucasian; step-wise screening process)	Consistent	Precise	Low

Depression: Screened vs. Unscreened					
1; 7380	RCT/ Fair: (High attrition rate)	Fair: (Participants higher risk than average screened population; Step-wise screening process)	N/A	Imprecise	Low
Depression: Screen Positive vs. Screen Negative					
3; 11,046	Prospective Cohort/ Fair to good: (One had high attrition, two had unclear discussion of masking)	Fair to good (one study likely had higher risk than the average screened population)	Consistent	Imprecise	Moderate
Negative Well-Being: Screen-detected vs. General Clinical Practice					
1; 165	Prospective Cohort/ Fair: (High attrition rates)	Fair: (Primarily Caucasian, Participants likely higher risk than average screened population; step-wise screening process)	N/A	Precise	Low
Negative Well-Being: Screened Positive vs. Screened Negative					
2; 319	Prospective Cohort/ Fair (High attrition)	Fair: (Primarily Caucasian, Participants likely higher risk than average screened population; step-wise screening process)	Consistent	Imprecise	Low

Cognitive Distress: Screen-detected vs. General Clinical Practice					
1; 165	Prospective Cohort/ Fair: (High attrition)	Fair: (Primarily Caucasian, Participants likely higher risk than average screened population; step-wise screening process)	N/A	Imprecise	Low
Cognitive Distress: Screened Positive vs. Screened Negative					
1; 319	Prospective Cohort/ Fair: (High attrition rates)	Fair: (Primarily Caucasian, Participants likely higher risk than average screened population; step-wise screening process)	N/A	Precise	Low
Worry: Screened vs. Unscreened					
1; 7380	RCT/ Fair: (High attrition rate)	Fair: (Participants likely higher risk than average screened population; Step-wise screening process)	N/A	Imprecise	Low
False Reassurance: Screened vs. Unscreened					
1; 5334	Prospective Cohort/ Fair: (High attrition rate)	Fair: (Participants likely higher risk than average screened population; Step-wise screening process)	N/A	Imprecise	Low

Measures of Cognition					
1; 431	Prospective Cohort/ Fair (Baseline differences not addressed in analysis)	Fair to good (primarily Caucasian, were identified through siblings with T2DM, which is unusual)	N/A	Precise	Low
KQ 2: Depression					
1; 3285	Prospective Cohort/Fair to good (unclear discussion of masking, missing data not addressed)	Good	N/A	Precise	Low

The evidence suggested that screening produces a significant increase in short-term anxiety up to 6 weeks later compared to no screening; however this anxiety does not persist at 3 months to 1 year after screening. Additionally, those who screen positive appear to experience higher anxiety immediately after diagnosis than those who screen negative, but again this difference is not present at 1 year. Interestingly, one cohort study showed that both people who screened positive and those who screened negative had lower anxiety scores at 1 year than before testing, indicating that perhaps the anticipation of screening initially made participants anxious.[27] This finding must be interpreted with caution, however, because it does not indicate whether anxiety measured prior to screening is anticipatory or reflects the participants' usual, baseline levels. The screening process also appeared to have a dose-response effect on anxiety, with those progressing further reporting higher, clinically-significant levels. Therefore, available evidence suggests that both the screening process itself as well as a subsequent diagnosis both result in temporary anxiety.

The short-term nature of anxiety found in this review fits with a previous systematic review, which showed that upon receiving a result that predicts higher risk of illness, anxiety, along with depression and distress, rose for only the first 4 weeks.[37] Although the strength of evidence is low for this outcome, the apparent dose-response relationship further supports the possibility that screening contributes to higher anxiety.

One trial showed that depression levels were similar between screened and unscreened groups. Additionally, 3 cohort studies showed no difference in depressive symptoms and 2 studies showed no difference in negative mood between those with screen-detected blood glucose abnormalities and those who screen negative. However, individuals with screen-detected diabetes who subsequently received treatment had higher odds of depression compared to those who screened negative. Although the latter comparison does not directly approximate the effects of treating screen-detected diabetes, it does support previous hypotheses suggesting that depression may be more related to treatment and its intensiveness than to the diagnosis of diabetes itself.[14,23,38] This is also important because only two of the four studies assessing depression reported how many patients with diabetes were taking medications; one reported low rates of medication usage, and the other reported the above association between treated T2DM and depression.[23,28] If the degree of diabetes management predicts incident depressive symptoms, a relatively low percentage of treatment among participants with screen-detected diabetes could explain a lack of difference in outcomes between these individuals and those who screened negative. Therefore, due to insufficient information about treatment, the available evidence in this review is insufficient to determine the full effect of the diabetes screening cascade on depression.

More people who screen positive for T2DM experience cognitive distress for up to 6 months following diagnosis than do those who screen negative. Screen-detected individuals also appear to have less distress at 2 weeks than those who are clinically detected; however this difference is very small and disappears at 6-12 months. Of note, these studies considered cognitive distress a symptom of diabetes; however it is not possible from the study design to determine whether this distress results from physiologic changes, the diagnosis, or the treatment of diabetes.

Diabetes-specific worry appears not to be increased in screened vs. unscreened individuals; however those with screen-detected diabetes have greater worry compared to those who screen negative. Therefore, the lack of difference between screened and unscreened groups may be due in part to the fact that a large number of screened people had negative results. Additionally, those progressing further through the screening process had greatest worry. As with anxiety, this apparent dose-response relationship suggests that screening for diabetes does increase worry, though this effect may be limited to those who screen positive. However, even among individuals receiving positive results, overall worry is low, which may be explained by psychological adjustment during the stepwise screening process to the possibility of having diabetes.

Despite low overall worry, those receiving results indicating that they are at risk for or currently have diabetes tend to think more about what it would be like to live with diabetes. Therefore, an "abnormal" result may increase thought intrusiveness, which could explain higher anxiety, distress or worry among those testing positive.

Limitations of the Methods

This review has several limitations. First, due to resources, a dual-review of inclusion, data abstraction, and strength of evidence rating was infeasible. Therefore, these results reflect one person's analysis and could change upon review by another analyst.

Second, only prospective studies were included, which may have limited the thoroughness of the analysis. However, even if a cross-sectional study were to show an association between screening and a psychological outcome, the very nature of the study design would limit the clarity of the causal relationship. In such an event, the next logical step would be to examine the association prospectively, as all of the included studies have done. Therefore, the likelihood that cross-sectional data would meaningfully change the results is low.

Finally, I did not include only those studies that examined the overall effect of screening by comparing screened vs. unscreened groups. For example, many studies measured outcome differences between people diagnosed with screen-detected diabetes and those who underwent screening but received a negative result; this comparison approximates only the effect of living with the diagnosis of diabetes as a result of screening. In contrast to the former comparison, these studies would not provide evidence for the effects of either being invited for screening or waiting for the screening test. Nevertheless, the effects of diagnosis and treatment as a result of screening are necessary to provide valuable evidence about the entire screening cascade.

Limitations of the Literature

The included studies contain some methodological issues that may mask a true difference in outcomes between groups, if one exists. For example, all studies had loss to follow-up, some to a much higher degree than others; however dropout rates are likely not at random, meaning

that the reasons for failing to submit a follow-up questionnaire may be related to the outcome of interest.[24,39] Indeed, prior evidence shows that likelihood of loss to follow-up is associated with the degree of anxiety and distress.[40,41] In other words, those who are more likely to have adverse psychological consequences following screening and/or diagnosis, may be less likely to follow-up with providers, which could differentially affect outcome data. Therefore, the difference between groups that drop out may be greater than the difference between groups that remain in the study.

Another limitation is that the studies generally are not able to assess how patients interact with their providers regarding diagnosis and treatment.[24] Because the way in which providers frame the potential seriousness of diabetes may affect patients' attitudes and perceptions, it represents an unmeasured determinant of psychological well-being. For example, if providers downplay the significance of an "at-risk" screening result, it could mitigate some of the negative psychological effects that screening otherwise would have produced.

The psychological outcomes included in this review can be divided into two major categories: general vs. disease-specific. Most of the studies in this review focus on general psychological harms such as depression, anxiety and quality of life. Unfortunately, these general quality of life measures may miss certain aspects unique to diabetes, which could partially explain the apparently small effect that screening has on psychological outcomes.[13] Indeed, the diabetes-specific measures such as worry, intrusiveness and cognitive distress tended to show a small but significant association with screening or with a diagnosis of screen-detected diabetes. Additionally, past longitudinal research shows that a disease specific outcome, diabetes-related distress, occurs more frequently among those with diabetes than does generalized anxiety or major depressive disorder.[42] Therefore, the relative lack of prospective studies that assess

diabetes-specific psychological outcomes means that current evidence may underestimate the harms of screening, or at the very least, may paint an incomplete picture of its overall psychological effects.

Implications

The overall strength of evidence for the outcomes in this systematic review is low; however the results have important implications for practice. The available evidence suggests that screening for type 2 diabetes may not be entirely benign; rather it may produce a small psychological harm in the form of increased anxiety, distress, worry, or thought intrusiveness. Although these harms may appear minimal, it is important to keep in mind that at best, the benefits of screening are also small. Therefore, even if the harms seem minor, they can meaningfully offset potential benefits such that the net benefit is too small to recommend routine screening. Unfortunately, the relative benefits and harms of screening are subject to a large degree of uncertainty, as gaps in the evidence still remain.

Future Research

In order to more fully estimate the potential harms of screening for T2DM, future research should focus on prospective studies that assess diabetes-related psychological harms. For example, studies may focus on diabetes-related distress, measured by validated scales such as the Problem Areas in Diabetes (PAID) scale. They may compare screened and unscreened individuals, as well as those who screen positive and screen negative to determine the effect of screening and diagnosis on diabetes-related distress. Following patients from the pre-screening period through diagnosis and treatment of screen-detected diabetes will provide a longitudinal account of how patients psychologically handle disease management. Finally, regular assessment

of intermediate outcomes such as hemoglobin A1C can offer information about how distress and diabetes management influence each other, that is, whether distress determines success of management or vice versa.

Conclusion

Overall, the available literature is insufficient to say with high certainty that screening for T2DM results in meaningful psychological harm, though the included studies hint that screening is not entirely benign. This review does provide weak evidence suggesting that screening for diabetes is associated with certain psychological harms. The screening process itself produces short-term anxiety and worry, whereas the diagnosis of screen-detected diabetes is associated with increased cognitive distress and diabetes-related thought intrusion. Treatment, but not screening and diagnosis, appears to be associated with elevated depressive symptoms. Unfortunately, the available evidence does not cover the entire spectrum of potential psychological harms. Only with future research that prospectively investigates more diabetes-specific psychological harms will we have a more accurate picture of the relative benefits and harms of diabetes screening.

References

1. Centers for Disease Control and Prevention. National diabetes fact sheet: National estimates and general information on diabetes and prediabetes in the united states. . 2011.

2. Norris S, Kansagara D, Bougatsos C, Fu R. Screening adults for type 2 diabetes: A review of the evidence for hte U.S. preventive services task force. *Annals of Internal Medicine*. 2008;148(11):855.

3. Echouffo-Tcheugui JB, Ali MK, Griffin SJ, Narayan KM. Screening for type 2 diabetes and dysglycemia. *Epidemiol Rev*. 2011;33(1):63-87. doi: 10.1093/epirev/mxq020; 10.1093/epirev/mxq020.

4. Harris R, Sawaya GF, Moyer VA, Calonge N. Reconsidering the criteria for evaluating proposed screening programs: Reflections from 4 current and former members of the U.S. preventive services task force. *Epidemiol Rev*. 2011;33(1):20-35. doi: 10.1093/epirev/mxr005; 10.1093/epirev/mxr005.

5. Simmons RK, Echouffo-Tcheugui JB, Sharp SJ, et al. Screening for type 2 diabetes and population mortality over 10 years (ADDITION-cambridge): A cluster-randomised controlled trial. *Lancet*. 2012;380(9855):1741-1748. doi: 10.1016/S0140-6736(12)61422-6; 10.1016/S0140-6736(12)61422-6.

6. American Diabetes Association. Standards of medical care in Diabetes—2013. *Diabetes Care*. 2013;36(Supplement 1):S11-S66. doi: 10.2337/dc13-S011.

7. Rosenzweig J, Ferrannini E, Grundy S, et al. Primary prevention of cardiovascular disease and type 2 diabetes in patients at metabolic risk: An endocrine society clinical practice guideline. *Journal of Clinical Endocrinology & Metabolism*. 2008;93(10):3671.

8. American Heart Association. Symptoms, diagnosis, and monitoring of diabetes.

9. United States Preventive Services Task Force. Screening for type 2 diabetes mellitus in adults, topic page.

10. National Institute for Health and Clinical Excellence. Preventing type 2 diabetes: Risk identification and interventions for individuals at high risk. . 2013:1-162.

11. Canadian Task Force on Preventive Health Care. Recommendations on screening for type 2 diabetes in adults. *Canadian Medical Association Journal*. 2012;184(16):1815-1815. doi: 10.1503/cmaj.112-2080.

12. International Diabetes Federation. About diabetes: Prevention.

13. Watkins K, Connell CM. Measurement of health-related QOL in diabetes mellitus. *Pharmacoeconomics*. 2004;22(17):1109-1126.

14. Thoolen BJ, de Ridder DT, Bensing JM, Gorter KJ, Rutten GE. Psychological outcomes of patients with screen-detected type 2 diabetes: The influence of time since diagnosis and treatment intensity. *Diabetes Care*. October 2006;29(10):2257-2262. doi: 10.2337/dc06-0617.

15. Polonsky WH, Anderson BJ, Lohrer PA, et al. Assessment of diabetes-related distress. *Diabetes Care*. 1995;18(6):754-760.

16. Harris RP, Helfand M, Woolf SH, et al. Current methods of the U.S. preventive services task force: A review of the process. *Am J Prev Med*. 2001;20(3, Supplement 1):21-35. doi: 10.1016/S0749-3797(01)00261-6.

17. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: Grading the strength of a body of evidence when comparing medical interventions--agency for healthcare research and quality and the effective health-care program. *J Clin Epidemiol*. 2010;63(5):513-523. doi: 10.1016/j.jclinepi.2009.03.009; 10.1016/j.jclinepi.2009.03.009.

18. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Ann Intern Med*. 2009;151(4):264-9, W64.

19. Paddison CA, Eborall HC, Sutton S, et al. Are people with negative diabetes screening tests falsely reassured? parallel group cohort study embedded in the ADDITION (cambridge) randomised controlled trial. *BMJ*. 2009;339:b4535. doi: 10.1136/bmj.b4535.

20. Park P, Simmons RK, Prevost AT, Griffin SJ. Screening for type 2 diabetes is feasible, acceptable, but associated with increased short-term anxiety: A randomised controlled trial in british general practice. *BMC Public Health*. 2008;8.

21. Eborall HC, Griffin SJ, Prevost AT, Kinmonth A-, French DP, Sutton S. Psychological impact of screening for type 2 diabetes: Controlled trial and comparative study embedded in the ADDITION (cambridge) randomised controlled trial. *Br Med J*. 2007;335(7618):486-489.

22. Adriaanse MC, Dekker JM, Spijkerman AMW, et al. Health-related quality of life in the first year following diagnosis of type 2 diabetes: Newly dignosed patients in general practice compared with screening-detected patients. the hoorn screening study. *Diabetic Med*. 2004;21(10):1075-1081.

23. Icks A, Albers B, Haastert B, et al. Risk for high depressive symptoms in diagnosed and previously undetected diabetes: 5-year follow-up results of the heinz nixdorf recall study. *PLoS ONE*. 2013;8(2):e56300. http://dx.doi.org/10.1371%2Fjournal.pone.0056300.

24. Adriaanse MC, Snoek FJ, Dekker JM, et al. No substantial psychological impact of the diagnosis of type 2 diabetes following targeted population screening: The hoorn screening study. *Diabetic Med*. 2004;21(9):992-998.

25. Adriaanse MC, Dekker JM, Spijkerman AMW, et al. Diabetes-related symptoms and negative mood in participants of a targeted population-screening program for type 2 diabetes: The hoorn screening study. *Qual Life Res*. 2005;14(6):1501-1509.

26. Edelman D, Olsen MK, Dudley TK, Harris AC, Oddone EZ. Impact of diabetes screening on quality of life. *Diabetes Care*. 2002;25(6):1022-1026.

27. Farmer AJ, Doll H, Levy JC, Salkovskis PM. The impact of screening for type 2 diabetes in siblings of patients with established diabetes. *Diabetic Med*. 2003;20(12):996-1004.

28. Golden SH, Lazo M, Carnethon M, et al. Examining a bidirectional association between depressive symptoms and diabetes. *JAMA*. 2008;299(23):2751-2759. doi: 10.1001/jama.299.23.2751; 10.1001/jama.299.23.2751.

29. Paddison CA, Eborall HC, French DP, et al. Predictors of anxiety and depression among people attending diabetes screening: A prospective cohort study embedded in the ADDITION (cambridge) randomized control trial. *Br J Health Psychol*. 2011;16(Pt 1):213-226. doi: 10.1348/135910710X495366; 10.1348/135910710X495366.

30. Rahman M, Simmons RK, Hennings SH, Wareham NJ, Griffin SJ. How much does screening bring forward the diagnosis of type 2 diabetes and reduce complications? twelve year follow-up of the ely cohort. *Diabetologia*. 2012;55(6):1651-1659.

31. Rahman M, Simmons RK, Hennings SH, Wareham NJ, Griffin SJ. Effect of screening for type 2 diabetes on population-level self-rated health outcomes and measures of cardiovascular risk: 13-year follow-up of the ely cohort. *Diabet Med*. 2012;29(7):886-892. doi: 10.1111/j.1464-5491.2012.03570.x; 10.1111/j.1464-5491.2012.03570.x.

32. Eborall H, Davies R, Kinmonth AL, Griffin S, Lawton J. Patients' experiences of screening for type 2 diabetes: Prospective qualitative study embedded in the ADDITION (cambridge) randomised controlled trial. *BMJ*. 2007;335(7618):490. doi: 10.1136/bmj.39308.392176.BE.

33. Roy T, Lloyd CE, Pouwer F, Holt RIG, Sartorius N. Screening tools used for measuring depression among people with type 1 and type 2 diabetes: A systematic review. *Diabet Med*. 2012;29(2):164-175.

34. Julian LJ. Measures of anxiety: State-trait anxiety inventory (STAI), beck anxiety inventory (BAI), and hospital anxiety and depression scale-anxiety (HADS-A). *Arthritis Care & Research*. 2011;63(S11):S467-S472. doi: 10.1002/acr.20561.

35. Smarr KL, Keefer AL. Measures of depression and depressive symptoms: Beck depression inventory-II (BDI-II), center for epidemiologic studies depression scale (CES-D), geriatric depression scale (GDS), hospital anxiety and depression scale (HADS), and patient health questionnaire-9 (PHQ-9). *Arthritis Care & Research*. 2011;63(S11):S454-S466. doi: 10.1002/acr.20556.

36. Lerman C, Trock B, Rimer BK, Jepson C, Brody D, Boyce A. Psychological side effects of breast cancer screening. *Health Psychol*. 1991;10(4):259-267.

37. Shaw C, Abrams K, Marteau TM. Psychological impact of predicting individuals' risks of illness: A systematic review. *Soc Sci Med*. 1999;49(12):1571-1598. doi: http://dx.doi.org.libproxy.lib.unc.edu/10.1016/S0277-9536(99)00244-0.

38. Nouwen A, Nefs G, Caramlau I, et al. Prevalence of depression in individuals with impaired glucose metabolism or undiagnosed diabetes: A systematic review and meta-analysis of the european depression in diabetes (EDID) research consortium. *Diabetes Care*. 2011;34(3):752-762. doi: 10.2337/dc10-1414; 10.2337/dc10-1414.

39. Stolk RP. Screening for diabetes. *Br Med J*. 2007;335(7618):457-458.

40. Lerman C, Daly M, Sands C, et al. Mammography adherence and psychological distress among women at risk for breast cancer. *J Natl Cancer Inst*. 1993;85(13):1074-1080.

41. Lerman C, Miller SM, Scarborough R, Hanjani P, Nolte S, Smith D. Adverse psychologic consequences of positive cYtologic cervical screening. *Obstet Gynecol*. 1991;165(3):658-662. doi: http://dx.doi.org.libproxy.lib.unc.edu/10.1016/0002-9378(91)90304-A.

42. Fisher L, Skaff MM, Mullan JT, Arean P, Glasgow R, Masharani U. A longitudinal study of affective and anxiety disorders, depressive affect and diabetes distress in adults with type 2 diabetes. *Diabetic Med*. 2008;25(9):1096-1101.

Embase: 4/3/2013

1. 'non insulin dependent diabetes mellitus'/exp

   AND

2. ('screening'/exp OR 'hemoglobin a1c'/exp)

   AND

3. ('depression'/exp OR 'stress'/exp OR 'distress syndrome'/exp OR 'emotion'/exp OR 'quality of life'/exp OR 'wellbeing'/exp)

   AND

4. [humans]/lim

   AND

5. [embase]/lim

Pubmed 4/5/2013

1. (diabetes mellitus, type 2 OR "type 2 diabetes" OR "non insulin dependent diabetes" OR "impaired fasting glucose" OR "impaired glucose tolerance")

   AND

2. (mass screening[MH] OR screen[tw] OR early diagnosis[tw] OR early detection[tw] OR "hemoglobin a1c"[tw] OR Hemoglobin A, Glycosylated[MH] OR "fasting glucose" OR "glucose tolerance test"[MH])

   AND

3. (depression[tw] OR depressed[tw] OR depressive[tw] OR distress[tw] OR stress[tw] OR stressed[tw] OR stressful[tw] OR worry[tw] OR worried[tw] OR fear*[tw] OR anxiety[tw] OR anxious[tw] OR quality of life[tw] OR mental health[tw] OR mental disorders[tw] OR psycholog*[tw] OR well being[tw] OR psychosocial[tw] OR uncertainty[tw] OR emotion*[tw])

Appendix A: Search Results

PsycInfo 4/5/2013:

1. T1 diabetes mellitus type 2 OR type 2 diabetes

   AND

2. ((((((DE "Major Depression") OR (DE "Anxiety")) OR (DE "Stress")) OR (DE "Distress")) OR (DE "Quality of Life")) OR (DE "Well Being")) OR (DE "Uncertainty")

   AND

3. DE "Screening Tests" OR DE "Screening"

Appendix B: Characteristics of Included Studies

Author, Year Study Name	Country	Study Design	Patients, n	Comparison Groups	Inclusion Criteria	Outcome	Time of Assessment
Rahman et al., 2012[30] How much does screening bring forward the diagnosis of type 2 diabetes and reduce complications? Twelve year follow-up of the Ely cohort	UK	Prospective Cohort	Total: 152 G1: 92 G2: 60	G1: People with T2DM screened in 1990-1992 G2: People with T2DM screened in 2000-2002	Adults age 40-65 living in Cambridgeshire without previously diagnosed T2DM attending a single practice	Mental Health	2002-2003 (10-12 years after G1 screening; 1-2 years after G2 screening)
Adriaanse et al., 2005[25] Diabetes-related symptoms and negative mood in participants of a targeted population-screening program for type 2 diabetes: The Hoorn Screening Study	The Netherlands	Prospective Cohort	Total: 319 G1:156 G2:163	G1: Screen-detected T2DM G2: Screen negative for T2DM	Adults 50-75 who were "high-risk" (>6 on Symptom Risk Questionnaire) were invited for screening	1: Psychological fatigue 2:Psychological cognitive distress 3: Negative mood	At 2 weeks, 6 months, and 12 months after diagnosis

Appendix B: Characteristics of Included Studies

Author, Year Study Name	Country	Study Design	Patients, n	Comparison Groups	Inclusion Criteria	Outcome	Time of Assessment
Adriaanse et al., 2003 Health-related quality of life in the first year following diagnosis of Type 2 diabetes: newly diagnosed patients in general practice compared with screening-detected patients. The Hoorn Screening Study	The Netherlands	Prospective Cohort	Total: 165 G1: 49 G2: 116	G1: Newly detected T2DM in general practice G2: Screen-detected T2DM	Adults 50-75 who were recently clinically diagnosed with T2DM. Adults 50-75 "high-risk" (>6 on Symptom Risk Questionnaire) were invited for screening	1:Psychological Fatigue 2: Cognitive Distress 3: Mental Health 4: Negative well-being	At 2 weeks, 6 months, and 12 months after diagnosis
Adriaanse et al., 2003 No substantial psychological impact of the diagnosis of Type 2 diabetes following targeted population screening: The Hoorn Screening Study	The Netherlands	Prospective Cohort	Total: 259 G1: 116 G2: 143	G1: Screen-detected T2DM G2: Screen negative for T2DM	Adults 50-75 who were "high-risk" (>6 on Symptom Risk Questionnaire) were invited for screening	1. Negative well-being 2. Mental Health	At 2 weeks, 6 months, and 12 months after diagnosis

Appendix B: Characteristics of Included Studies

Author, Year Study Name	Country	Study Design	Patients, n	Comparison Groups	Inclusion Criteria	Outcome	Time of Assessment
Farmer et al., 2003 The impact of screening for Type 2 diabetes in siblings of patients with established diabetes	UK	Prospective Cohort	Total:431 G1: 227 G2: 101 G3: 85 G4: 18	G1: Normoglycemic G2: Borderline G3: High Risk G4: Possible Diabetes	Adults 35-74 years without diabetes who have a sibling with T2DM and were identified as willing to participate by sibling	Primary: 1. Anxiety Secondary: 2. Perceived risk of diabetes 3. Extent to which thoughts were occupied with diabetes/complications	Primary: Before screening and 1 year post-diagnosis Secondary: At screening and 1 year post-diagnosis
Rahman et al., 2012 How much does screening bring forward the diagnosis of type 2 diabetes and reduce complications? Twelve year follow-up of the Ely cohort	UK	Prospective Cohort	Total: 1441 G1: 731 G2: 711	G1: People screened negative for T2DM in 1990-1992 G2: People screened negative for T2DM in 2000-2003	Adults age 40-65 living in Cambridgeshire without previously diagnosed T2DM who attend a single practice	Mental Health	2002-2003 (10-13 years after G1 screening; 1-3 years after G2 screening)

Appendix B: Characteristics of Included Studies

Author, Year Study Name	**Country**	**Study Design**	**Patients, n**	**Comparison Groups**	**Inclusion Criteria**	**Outcome**	**Time of Assessment**
Paddison et al., 2011 Predictors of anxiety and depression among people attending diabetes screening:A propspective cohort study embedded in the ADDITION (Cambridge) randomized controlled trial	UK	Prospective Cohort	Total: 3240	G1: Screen-detected T2DM G2: Screen negative for T2DM	Adults 40-69 years in top quartile of Cambridge Diabetes Risk Score were invited to screening	1.Anxiety 2.Depression	Following known results of RBG and at 12 months follow-up
Paddison et al., 2009 Are people with negative diabetes screening tests falsely reassured? Parallel group cohort study embedded in the ADDITION (Cambridge) randomised controlled trial	UK	Prospective Cohort	Total: 5334 G1:964 G2: 4370	G1: Unscreened G2: Screened negative for T2DM	Adults 40-69 years in top quartile of Cambridge Diabetes Risk Score were invited to screening or questionnaire without screening	False reassurance 1. Personal Risk 2. Comparative Risk	After initial screening test, at 3-6 months, then 12-15 months after screening

Appendix B: Characteristics of Included Studies

Author, Year Study Name	**Country**	**Study Design**	**Patients, n**	**Comparison Groups**	**Inclusion Criteria**	**Outcome**	**Time of Assessment**
Golden et al., 2008 Examining a Bidirectional Association Between Depressive Symptoms and Diabetes (Multi-Ethnic Study of Atherosclerosis)	US	Prospective Cohort	Total: 4847 G1: 2868 G2: 1357 G3: 203 G4: 417	G1: Screen negative for T2DM G2: Screened Impaired Fasting Glucose G3: Screened positive for T2DM-untreated G4: Screened positive for T2DM-treated	Men and women 45-84 white, black, Hispanic, Chinese without self-reported cardiovascular disease without baseline depressive symptoms or antidepressant medication use	Depressive symptoms	5 year follow-up
Edelman et al., 2002 Impact of Diabetes Screening on Quality of Life	US	Prospective Cohort	Total:1253 G1:1177 G2: 56	G1: Screened-negative for T2DM G2: Screened-positive for T2DM	Men and women 45-64 at outpatient DVAMC without diabetes at baseline	Mental Health	Baseline and 1 year after enrollment

Appendix B: Characteristics of Included Studies

Author, Year Study Name	**Country**	**Study Design**	**Patients, n**	**Comparison Groups**	**Inclusion Criteria**	**Outcome**	**Time of Assessment**
Icks et al, 2013 Risk for High Depressive Symptoms in Diagnosed and Previously Undetected Diabetes: 5-year follow-up results of the Heinz Nixdorf Recall Study	Germany	Prospective Cohort	Total 3633	G1: Previously Diagnosed T2DM G2: Screen-detected T2DM G3: Screen-negative for T2DM	Men and women 45-75 without depressive symptoms at baseline	Depressive symptoms	At 5 years
Park et al., 2008 Screening for type 2 diabetes is feasible, acceptable, but associated with increased short-term anxiety: A randomized controlled trial in British general practice	UK	Randomized Controlled Trial	Total: 355 G1: 238 G2: 116	Individual-randomization G1: Unscreened G2: Screened	Adults 40-69 without known diabetes, identified as high risk using a risk score	1. Anxiety	6 weeks post screening or invitation

Appendix B: Characteristics of Included Studies

Author, Year Study Name	Country	Study Design	Patients, n	Comparison Groups	Inclusion Criteria	Outcome	Time of Assessment
Eborall et al., 2008 Psychological impact of screening for type 2 diabetes: controlled trial and comparative study embedded in the ADDITION (Cambridge) randomised trial	UK	Non-randomized controlled trial	Total: 7380 G1: 6416 G2: 964	G1: Screened G2: Unscreened	Adults 40-69 in the top fourth for risk of haivng undiagnosed type 2 diabetes	1.State Anxiety 2. Anxiety 3. Depression 4. Worry	At RBG, 3-6 months, and 12-15 months

Appendix C: Internal Validity Ratings

Author, Year Study Name	**Was randomization adequate?** **Was allocation concealment adequate?**	**Were groups similar at baseline?**	**Were outcome assessors/ providers/ patients masked?**	**Was overall attrition ≥20%?**	**Was differential attrition ≥15%?**	**Did the study use ITT analyses?**	**Proper analysis? (adjustment for confounders)**	**Were outcome measures equal, valid and reliable?**	**Overall Internal Validity**
Rahman et al., 2012 How much does screening bring forward the diagnosis of type 2 diabetes and reduce complications? Twelve year follow-up of the Ely cohort	N/A-Cohort Study N/A-Cohort Study	No: Recently screened group had more people on antiplatelet drugs, had higher retinopathy rates, and had more people with ECG confirmed ischemic heart disease compared to screened in the past	Yes/No/No	Yes	No	N/A-Cohort Study	No-did not adjust for baseline differences, did not adjust for multiple testing	Yes	Fair

Appendix C: Internal Validity Ratings

Author, Year Study Name	**Was randomization adequate?** **Was allocation concealment adequate?**	**Were groups similar at baseline?**	**Were outcome assessors/ providers/ patients masked?**	**Was overall attrition ≥20%?**	**Was differential attrition ≥15%?**	**Did the study use ITT analyses?**	**Proper analysis? (adjustment for confounders)**	**Were outcome measures equal, valid and reliable?**	**Overall Internal Validity**
Adriaanse et al., 2005 Diabetes-related symptoms and negative mood in participants of a targeted population-screening program for type 2 diabetes: The Hoorn Screening Study	N/A-Cohort Study N/A-Cohort Study	No: Screen positive group had higher blood glucose and slightly higher symptom risk score	N/A-Self-reported questionnaire by mail/ No /No	Yes	No	N/A-Cohort Study	No-did not adjust for baseline differences, did not account for multiple testing	Yes	Fair

Appendix C: Internal Validity Ratings

Author, Year Study Name	**Was randomization adequate?** **Was allocation concealment adequate?**	**Were groups similar at baseline?**	**Were outcome assessors/ providers/ patients masked?**	**Was overall attrition ≥20%?**	**Was differential attrition ≥15%?**	**Did the study use ITT analyses?**	**Proper analysis? (adjustment for confounders)**	**Were outcome measures equal, valid and reliable?**	**Overall Internal Validity**
Adriaanse et al., 2003 Health-related quality of life in the first year following diagnosis of Type 2 diabetes: newly diagnosed patients in general practice compared with screening-detected patients. The Hoorn Screening Study	N/A-Cohort Study N/A-Cohort Study	No: Screen-detected group had lower A1c, were more overweight, were more likely hypertensive, much less likely to be on blood glucose lowering medications,	N/A-Self-reported questionnaire by mail / No /No	Yes	No	N/A-Cohort Study	No-Did not adjust for baseline differences, did not account for multiple testing, which would have made everything non-significant	Yes	Fair

Appendix C: Internal Validity Ratings

Author, Year Study Name	**Was randomization adequate?** **Was allocation concealment adequate?**	**Were groups similar at baseline?**	**Were outcome assessors/ providers/ patients masked?**	**Was overall attrition ≥20%?**	**Was differential attrition ≥15%?**	**Did the study use ITT analyses?**	**Proper analysis? (adjustment for confounders)**	**Were outcome measures equal, valid and reliable?**	**Overall Internal Validity**
Adriaanse et al., 2003 No substantial psychological impact of the diagnosis of Type 2 diabetes following targeted population screening: The Hoorn Screening Study	N/A-Cohort Study N/A-Cohort Study	No-diabetes group had higher glucose levels, greater degree of dyspnea on exertion and less likely to use bicycle for transportation	N/A-Self-reported questionnaire by mail/ No/ No	No	No	N/A-Cohort Study	No-only adjusted for baseline scores. Did not adjust for other baseline imbalances and did not account for multiple testing	Yes	Fair

Appendix C: Internal Validity Ratings

Author, Year Study Name	**Was randomization adequate?** **Was allocation concealment adequate?**	**Were groups similar at baseline?**	**Were outcome assessors/ providers/ patients masked?**	**Was overall attrition ≥20%?**	**Was differential attrition ≥15%?**	**Did the study use ITT analyses?**	**Proper analysis? (adjustment for confounders)**	**Were outcome measures equal, valid and reliable?**	**Overall Internal Validity**
Farmer et al., 2003 The impact of screening for Type 2 diabetes in siblings of patients with established diabetes	N/A-Cohort Study N/A-Cohort Study	No- Those with diabetes were more likely male, had higher BMI; other characteristics unclear due to reporting	N/A-Self-reported questionnaire by mail/ No/ No	No	No	N/A-Cohort Study	No-did not adjust for past treatment of depression or any other potential confounders in list of baseline characteristics	Yes	Fair

Appendix C: Internal Validity Ratings

Author, Year Study Name	**Was randomization adequate?** **Was allocation concealment adequate?**	**Were groups similar at baseline?**	**Were outcome assessors/ providers/ patients masked?**	**Was overall attrition ≥20%?**	**Was differential attrition ≥15%?**	**Did the study use ITT analyses?**	**Proper analysis? (adjustment for confounders)**	**Were outcome measures equal, valid and reliable?**	**Overall Internal Validity**
Rahman et al., 2012 How much does screening bring forward the diagnosis of type 2 diabetes and reduce complications? Twelve year follow-up of the Ely cohort	N/A-Cohort Study N/A-Cohort Study	Yes	Yes/No/No	Yes (54 vs 46)	No	N/A-Cohort Study	No-there was no adjustment: the data were exploratory, but did not account for multiple testing	Yes	Fair: high attrition rate

Appendix C: Internal Validity Ratings

Author, Year Study Name	**Was randomization adequate?** **Was allocation concealment adequate?**	**Were groups similar at baseline?**	**Were outcome assessors/ providers/ patients masked?**	**Was overall attrition ≥20%?**	**Was differential attrition ≥15%?**	**Did the study use ITT analyses?**	**Proper analysis? (adjustment for confounders)**	**Were outcome measures equal, valid and reliable?**	**Overall Internal Validity**
Paddison et al., 2011 Predictors of anxiety and depression among people attending diabetes screening:A propspective cohort study embedded in the ADDITION (Cambridge) RCT	N/A-Cohort Study N/A-Cohort Study	Unclear/Not-reported	Unclear / No/No	Yes	Unclear/not-reported	N/A-Cohort Study	Yes in logistic regression, but not in chi square analysis of prevalence of depressive symptoms. Did not address multiple comparisons	Yes	Fair: high attrition rate

Appendix C: Internal Validity Ratings

Author, Year Study Name	**Was randomization adequate?** **Was allocation concealment adequate?**	**Were groups similar at baseline?**	**Were outcome assessors/ providers/ patients masked?**	**Was overall attrition ≥20%?**	**Was differential attrition ≥15%?**	**Did the study use ITT analyses?**	**Proper analysis? (adjustment for confounders)**	**Were outcome measures equal, valid and reliable?**	**Overall Internal Validity**
Paddison et al., 2009 Are people with negative diabetes screening tests falsely reassured? Parallel group cohort study embedded in the ADDITION (Cambridge) randomised controlled trial	N/A-Cohort Study N/A-Cohort Study	Yes	Unclear / No/No	Yes	Yes	N/A-Cohort Study	Yes-No measured confounders, but did not measure educational status (could be associated with understanding). Did not address missing data or multiple testing.	Measures were equal.Unclear as to validity and reliability	Fair

Appendix C: Internal Validity Ratings

Author, Year Study Name	**Was randomization adequate?** **Was allocation concealment adequate?**	**Were groups similar at baseline?**	**Were outcome assessors/ providers/ patients masked?**	**Was overall attrition ≥20%?**	**Was differential attrition ≥15%?**	**Did the study use ITT analyses?**	**Proper analysis? (adjustment for confounders)**	**Were outcome measures equal, valid and reliable?**	**Overall Internal Validity**
Golden et al., 2008 Examining a Bidirectional Association Between Depressive Symptoms and Diabetes	N/A-Cohort Study N/A-Cohort Study	No- Those with diabetes were more likely male, african american, were less educated, had lower income, and had higher rates of hypertension	Unclear/ No/No	No	Unclear/not-reported	N/A-Cohort Study	Yes-Although did not address missing data	Yes	Fair to good
Edelman et al., 2002 Impact of Diabetes Screening on Quality of Life	N/A-Cohort Study N/A-Cohort Study	Unclear/Not-reported	Unclear / No/No	No	No	N/A-Cohort Study	Yes-Although did not report baseline covariates, they used ANCOVA for measured baseline scores	Yes	Fair

Appendix C: Internal Validity Ratings

Author, Year Study Name	**Was randomization adequate?** **Was allocation concealment adequate?**	**Were groups similar at baseline?**	**Were outcome assessors/ providers/ patients masked?**	**Was overall attrition ≥20%?**	**Was differential attrition ≥15%?**	**Did the study use ITT analyses?**	**Proper analysis? (adjustment for confounders)**	**Were outcome measures equal, valid and reliable?**	**Overall Internal Validity**
Icks et al., 2013 Risk for High Depressive Symptoms in Diagnosed and Previously Undetected Diabetes: 5-year follow-up results of the Heinz Nixdorf Recall Study	N/A-Cohort Study N/A-Cohort Study	No-No diabetes group had more women, more exercise, lower BMI	Unclear/ Unclear/ No	No	No	N/A-Cohort Study	Yes, multiple logistic regression models adjusted for measured baseline differences; however did not measure steroid use	Yes	Fair to good
Park et al., 2008 Screening for type 2 diabetes is feasible, acceptable, but associated with increased short-term anxiety: A RCT in British general practice	Unclear/not reported: did not discuss randomization protocol Unclear: likely was given computerized randomization	Yes	Unclear/ No/No	Yes	No	No-analyzed completers only	Yes-No baseline indifferences, but did not assess baseline anxiety	Yes	Fair

Appendix C: Internal Validity Ratings

Author, Year Study Name	**Was randomization adequate?** **Was allocation concealment adequate?**	**Were groups similar at baseline?**	**Were outcome assessors/ providers/ patients masked?**	**Was overall attrition ≥20%?**	**Was differential attrition ≥15%?**	**Did the study use ITT analyses?**	**Proper analysis? (adjustment for confounders)**	**Were outcome measures equal, valid and reliable?**	**Overall Internal Validity**
Eborall et al., 2008 Psychological impact of screening for type 2 diabetes: controlled trial and comparative study embedded in the ADDITION (Cambridge) randomised trial	N/A-Initially randomized at practice level, but this substudy was not randomized in which clinics they used No-investigators knew which practices were coming from initial study group	Yes	Unclear/ No/No	Yes	Yes	No-analyzed completers only	Yes, though did not adjust for multiple comparisons	Yes	Fair

Appendix D: External Validity Ratings

Author, Year **Study Name**	**Population** **(G/F/P)**	**Situation** **(G/F/P)**	**Providers** **(G/F/P)**	**Overall External Validity**
Rahman et al., 2012 How much does screening bring forward the diagnosis of type 2 diabetes and reduce complications? Twelve year follow-up of the Ely cohort	Fair-All of the people in the study had diagnosed diabetes; people with screen negative for diabetes were excluded; almost all caucasian	Good	Good	Fair
Adriaanse et al., 2005 Diabetes-related symptoms and negative mood in participants of a targeted population-screening program for type 2 diabetes: The Hoorn Screening Study	Fair: Almost all caucasian, were identified through symptom questionnaire: may be unlike an asymptomatic screening group	Fair: step-wise screening process	Good	Fair
Adriaanse et al., 2003 Health-related quality of life in the first year following diagnosis of Type 2 diabetes: newly diagnosed patients in general practice compared with screening-detected patients. The Hoorn Screening Study	Fair: Almost all caucasian, were identified through symptom questionnaire: may be unlike an asymptomatic screening group	Fair: step-wise screening process	Good	Fair

Appendix D: External Validity Ratings

Author, Year **Study Name**	**Population** **(G/F/P)**	**Situation** **(G/F/P)**	**Providers** **(G/F/P)**	**Overall External Validity**
Adriaanse et al., 2003 No substantial psychological impact of the diagnosis of Type 2 diabetes following targeted population screening: The Hoorn Screening Study	Fair: Almost all caucasian, were identified through symptom questionnaire: may be unlike an asymptomatic screening group	Fair-Stepwise screening approach	Good	Fair
Farmer et al., 2003 The impact of screening for Type 2 diabetes in siblings of patients with established diabetes	Fair-People belonged to big families, and were primarily caucasian	Fair-People likely not used to being approached for screening on account of a family member screening positive	Good	Fair
Rahman et al., 2012 How much does screening bring forward the diagnosis of type 2 diabetes and reduce complications? Twelve year follow-up of the Ely cohort	Poor-None of the people in the study had diagnosed diabetes; they were excluded from the analysis; people were generally healthy and almost all caucasian	Good	Good	Fair
Paddison et al., 2011 Predictors of anxiety and depression among people attending diabetes screening:A propspective cohort study embedded in the ADDITION (Cambridge) randomized controlled trial	Fair: Patients were at "highest risk" for T2DM, likely higher than average screened person in U.S.; Affluent patients	Fair to poor: had step-wise screening process	Good	Fair

Appendix D: External Validity Ratings

Author, Year **Study Name**	**Population** **(G/F/P)**	**Situation** **(G/F/P)**	**Providers** **(G/F/P)**	**Overall External Validity**
Paddison et al., 2009 Are people with negative diabetes screening tests falsely reassured? Parallel group cohort study embedded in the ADDITION (Cambridge) randomised controlled trial	Fair: Patients were at "highest risk" for T2DM, likely higher than average screened person in U.S.; Affluent patients	Fair to poor: had step-wise screening process	Good	Fair
Golden et al., 2008 Examining a Bidirectional Association Between Depressive Symptoms and Diabetes	Good	Good	Good	Good
Edelman et al., 2002 Impact of Diabetes Screening on Quality of Life	Good	Good	Good	Good
Icks et al, 2013	Good	Fair: low percentage of screen-detected T2DM untreated	Good	Fair

Appendix D: External Validity Ratings

Author, Year **Study Name**	**Population** **(G/F/P)**	**Situation** **(G/F/P)**	**Providers** **(G/F/P)**	**Overall External Validity**
Park et al., 2008 Screening for type 2 diabetes is feasible, acceptable, but associated with increased short-term anxiety: A randomized controlled trial in British general practice	Fair: Patients were at "highest risk" for T2DM, likely higher than average screened person in U.S.; Affluent patients	Fair: step-wise screening process	Good	Fair
Eborall et al., 2008 Psychological impact of screening for type 2 diabetes: controlled trial and comparative study embedded in the ADDITION (Cambridge) randomised trial	Fair: Patients were at "highest risk" for T2DM, likely higher than average screened person in U.S.	Fair to poor: had step-wise screening process	Good	Fair